Appendix

Given is a data set of monthly data. We denote soil moisture by w, and T, P, E, R and G are temperature, precipitation, evaporation, runoff and recharge to groundwater respectively. The data set X(s,j,m) denotes any of the above as a function of space, year (1, N) and month. Given is an initial condition $X^{IC}(s,j_0,m)$, for example the most recent map, where j_0 is outside the range j=1..N. A suitable monthly climatology is removed from the data - henceforth X shall be the anomaly. A constructed analogue is defined as:

$$X^{CA}(s, j_0, m) = \sum_{j=1}^{N} \alpha_j \; X(s, j, m) \quad (1)$$

where the coefficents α are determined so as to minimize the difference between $X^{CA}(s, j_0, m)$ and $X^{IC}(s, j_0, m)$. The solution to this problem is given in Van den Dool(1994). Eq (1) is only a diagnostic statement. We now seek a forecast of variable Y (which could be soil moisture itself) as follows:

$$Y^{F} (s, j_0, m+\tau) = \sum_{j=1}^{N} \alpha_j \; Y(s, j, m+\tau) \quad (2)$$

How well does this work? In case X=Y=w, one can verify the soil moisture forecast, i.e. the left hand of (2), against the observed $w(s, j_0, m+\tau)$. In case X=w, and Y=T or P or E, one can verify temperature, precipitation or evaporation forecasts against observations. (Keep in mind that w and E 'observation' are calculated in an LDAS like scheme.) In principle one can construct an analogue on any initial multi-field state, but we found initial w to be the best for forecasting itself and forecasting the other fields! This is a testimony that soil moisture is probably the key as has been suspected by many for decades.

H. M. van den Dool, 1994: Searching for analogues, how long must one wait? Tellus, 46A, 314-324.

J. Huang , H. M. van den Dool and K. G. Georgakakos, 1996: Analysis of model-calculated soil moisture over the US (1931-1993) and applications to long range temperature forecasts. J Climate., 9, 1350-1362.